

FACULTY OF ENGINEERING AND
INFORMATION TECHNOLOGY

Novel Models in Recommender Systems
and Group Recommender Systems
for Improving Recommendations

Jorge Castro

A thesis submitted for the Degree of
Doctor of Philosophy

University of Technology Sydney

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I, Jorge Castro declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

This thesis is the result of a Collaborative Doctoral Research Degree program with University of Granada.

This research is supported by the Australian Government Research Training Program.

Signature of Candidate:

Production Note:

Signature removed prior to publication.

Date: 26th February 2019

ACKNOWLEDGEMENTS

Given that research work requires a huge effort and work, I want to dedicate a few lines to those who have been with me one way or another in my last years of work supporting me. Without them, I would not reach this goal.

Among them, I want to highlight my family. My parents Ángel and María Dolores, without their education and discipline I would not be able to aspire to such a tough task. To my wife, Isabel, because of her endless support, patience and company along my academic stage. To my brother, Ángel Darío, for the shared moments. To the rest of my family and friends who have contributed somehow in my doctoral stage.

In the academia, I want to thank my supervisors, Jie Lu and Guangquan Zhang for their trust and patience during these years and the opportunity given to me for joining them in the UTS, which has led me to present this dissertation. Regarding the Spanish academia, I want to thank to Luis Martínez, for his patience, trust and effort. He was the first person trusting my research skills. Special mention to Oscar Cerdón, Paco Herrera, Salvador García, Rosa M. Rodríguez, Manuel J. Barranco, Macarena Espinilla and Pedro Sánchez.

Also in the academia, I want to thank my peers during my stage as PhD student, to Hongshu Chen, Chenlian Hu and Brenda Wang. And of course, to all the student and colleagues of the DeSI group for the treatment and hospitality received.

Regarding my Spanish colleagues in the academia, I want to mention Manu Parra, Rosa Rodríguez, Sara, Pablo, Dani Peralta, Manuel Titos, Lala, Jose, Antonio, Edu, Fran, Juanan, Rafa, Sergio Gonzalez, Jesús Maillo, Elena, Jose Ángel, Eva, Nacho, Andrea, Salva, Julián Luengo, Eugenio, Francisco J. Estrella, Francisco J. Quesada, Raciél Yera, Yeleny Zulueta, Hong Bing, Jun Liu, Shuwei Chen and Francisco Moya, among others. Thanks.

And last but not least, I want to thank to my closest friends Sergio, Ana, Javi and Raúl, among others. Their support in relaxation moments to disconnect from work meant a lot to me.

Thanks to all for your support.

ABSTRACT

This thesis focuses on the improvement of recommendations within personalization processes applied to support users overcome the information overload problem. This thesis focuses on two aspects: group recommendation and context-aware recommendation. The first block of the thesis proposes four group recommendation models to overcome each of the following limitations of previous techniques: (i) loss of information and diversity due to the aggregation of ratings in a group profile, (ii) lack of techniques that consider the changes in users' behavior when gathering in groups through consensus reaching processes, (iii) lack of techniques that consider the influences among members' preferences in group recommendation, and (iv) lack of natural noise management techniques for group recommender systems. The second block of the thesis focuses on the integration of contextual information in individual and group recommendation models. Within this block, a context-aware recommendation model for the recommendation in the question answering domain is proposed, which considers the collaborative trend interest as the recommendation context. The second model within this block integrates context-aware and group recommendation extending a consensus-driven group recommendation model.

Table of Contents

1	Introduction	23
1.1	Motivation	23
1.2	Research questions	25
1.3	Research objectives	26
1.4	Research significance	30
1.4.1	Theoretical significance	30
1.4.2	Practical significance	32
1.5	Research methodology and process	33
1.5.1	Research methodology	33
1.5.2	Research process	33
1.6	Thesis structure	34
1.7	Publications related to this thesis	35
2	Literature Review	37
2.1	Introduction to background	37
2.2	Recommender systems	42
2.2.1	Research trends in recommender systems	55
2.3	Group recommender systems	56
2.4	Context-aware recommender systems	63
2.4.1	Contextual modeling approaches	66
2.4.2	Frameworks for context-aware recommender systems	69

2.4.3	Context-awareness in group recommender systems	69
2.5	Concepts used in the proposed methods	74
2.5.1	Hesitant Fuzzy sets	74
2.5.2	Group decision making and consensus reaching processes	78
2.5.3	Opinion dynamics	83
2.5.4	Natural noise	90
3	Hesitant Fuzzy Sets-based Group Recommender System	95
3.1	Introduction	95
3.2	Hesitant Group Recommender Model	96
3.2.1	Hesitant Modeling of group's and user's preferences . . .	98
3.2.2	Neighborhood formation with HPCC	99
3.2.3	Rating prediction	100
3.2.4	Group recommendation	101
3.3	Experimentation and evaluation	101
3.3.1	Techniques compared	102
3.3.2	Data set	102
3.3.3	Evaluation measures	103
3.3.4	Experiment results	104
3.3.5	Computational complexity	109
3.4	Summary	110
4	Consensus-driven Group Recommender Systems	111
4.1	Introduction	111
4.2	Consensus-driven Group Recommender System	113
4.2.1	Recommendation Phase	113
4.2.2	Consensus Phase	115
4.3	Experimentation and evaluation	120
4.3.1	Techniques compared	121

4.3.2	Data set	122
4.3.3	Evaluation measures	123
4.3.4	Experiment 1: group recommender system with the minimum cost consensus model	124
4.3.5	Experiment 2: group recommender system with the automatic consensus support system model based on feedback	125
4.3.6	Graphical Visualization	126
4.3.7	Computational complexity	129
4.4	Summary	130
5	Opinion Dynamics-based Group Recommender Systems	133
5.1	Introduction	133
5.2	Framework for group recommendation based on opinion dynamics	135
5.2.1	Group recommender system based on opinion dynamics (Pre-GROD)	137
5.2.2	Group recommendation based on opinion dynamics with consensus (GROD)	138
5.3	Experimentation and evaluation	141
5.3.1	Techniques compared	141
5.3.2	Datasets	142
5.3.3	Evaluation measures	143
5.3.4	Experiment 1: Pre-GROD in the general case	144
5.3.5	Experiment 2: GROD in groups without consensus	147
5.3.6	Computational complexity	150
5.4	Sensitivity analysis	151
5.4.1	Member elimination	154
5.4.2	Rating variation	157
5.5	Summary	159

6	Natural noise management in group recommender systems	161
6.1	Introduction	161
6.2	Natural noise management in group recommendation	163
6.2.1	Local natural noise management based on local information (NNM-LL)	164
6.2.2	Local natural noise management based on global information (NNM-LG)	166
6.2.3	Global natural noise management (NNM-GG)	167
6.2.4	Hybrid global-local natural noise management (NNM-H)	168
6.2.5	Illustrative example	169
6.3	Experimentation and evaluation	170
6.3.1	Techniques compared	171
6.3.2	Data set	172
6.3.3	Experiment results for recommendation aggregation-based GRS	173
6.3.4	Results in rating aggregation-based GRS	175
6.3.5	Discussion	177
6.3.6	Complexity and deployment	182
6.4	Summary	184
7	Context-aware question answering recommendation with semantic model	187
7.1	Introduction	187
7.2	Context-aware content-based recommender system for questions based on topic detection in current trend interest	190
7.2.1	QA domain semantic analysis	191
7.2.2	Building user's preference profile	192
7.2.3	Context model building	193
7.2.4	Users' profile contextualization	194

7.2.5	Prediction	195
7.3	Experimentation and evaluation.	195
7.3.1	Experimental procedure	196
7.3.2	Techniques compared	196
7.3.3	Datasets	197
7.3.4	Evaluation Measures	198
7.3.5	Results	199
7.4	Summary	203
8	REJA: Location-aware Consensus-driven Recommender System for Restaurants	205
8.1	Introduction	205
8.2	Restaurants of Jaén Recommender System: REJA	206
8.3	Location and trajectory-awareness for recommendations on the move.	208
8.3.1	Prototype: LT-REJA	212
8.4	Location-aware consensus-driven group recommendation	215
8.4.1	Prototype: CLG-REJA	219
8.5	Summary	222
9	Conclusions and further study	225
9.1	Conclusions	225
9.2	Further study	229
A	Searching groups without consensus.	269
B	Similarity measures	275

List of Figures

1.1	Structure and contents of this thesis.	35
2.1	User-based collaborative filtering scheme.	46
2.2	Scheme of the factorization of the rating matrix.	48
2.3	Decomposition of TFIDF matrix with singular Value decomposition. Note that s is a diagonal matrix with the singular values sorted in descending order.	51
2.4	The framework of group recommender system based on rating aggregation.	59
2.5	The framework of group recommender system based on recommendation aggregation.	59
2.6	Different approaches to integrate the contextual information in the recommendation [6].	65
2.7	Example of hesitant valuations	76
2.8	Group decision making.	79
2.9	Scheme of resolution of a group decision making problem with a consensus reaching process.	81
2.10	General CRP scheme	82
2.11	Opinion evolution for the weighting matrix shown in Eq. 2.27. . .	85
2.12	Example of similarity matrix of a group expressed as a graph. . . .	87
2.13	Opinion evolution for the weighting matrix shown in Eq. 2.31. . .	89

2.14	General scheme of natural noise management for individuals [242].	92
2.15	Classification of the ratings of natural noise management for individuals [242].	93
3.1	General scheme of HGRM	97
3.2	Results for $NRMSE \cdot \alpha + ILS \cdot (1 - \alpha)$	108
4.1	General scheme of the consensus-driven group recommender system	114
4.2	Detail of the recommendation phase of the consensus-driven group recommender system.	114
4.3	Consensus phase with the automatic consensus support system model based on feedback.	117
4.4	Results of area under receiver operating characteristic curve (AUC) for evaluated configurations.	125
4.5	Precision at certain recommendation list sizes for evaluated configurations.	126
4.6	Visualization of the recommendations made by the baseline GRS and the consensus-based GRS.	131
5.1	Framework for group recommender system based on opinion dynamics.	136
5.2	Group recommender system based on opinion dynamics with consensus (GROD).	140
5.3	Results for the evaluated GRSs in MovieLens 100k.	145
5.4	Results for the evaluated GRSs in MovieLens 1M.	146
5.5	Results for the evaluated GRSs in MovieLens-100k.	148
5.6	Results for the evaluated GRSs in MovieLens-1M.	150
5.7	Changes in the item ranking for group [126, 595, 607, 619, 648] and when user 126 leaves.	153

5.8	Results of Intersection@Size for groups of size 2 to 10 to over the recommendation list.	155
5.9	Results of Spearman@Size for groups of size 2 to 10 to over the recommendation list.	155
5.10	Changes in the recommendation when member leaves the group sorted by number of ratings.	156
5.11	Changes in the recommendation when member changes ratings.	158
6.1	Scheme of local NNM for GRS, showing two approaches: a) NNM-LL, b) NNM-LG.	165
6.2	Global natural noise management based on global information (NNM-GG) application.	167
6.3	Hybrid global-local natural noise management (NNM-H) application to group recommendation.	168
7.1	General scheme of the proposed context-aware content-based recommender system for questions based on topic detection in current trend interest.	191
7.2	Context model building phase.	193
7.3	Contextual dataset used in the experiment, where each day has a different status update count.	199
7.4	Results managing context without clustering	200
7.5	Results managing context with fuzzy membership.	201
7.6	Results managing context with max membership.	202
7.7	Results of the proposal to manage context with fuzzy membership.	203
7.8	Average NDCG of the compared approaches in all contexts.	204
8.1	Province of Jaén, location of interest of REJA	207

8.2	Displacing the AOI according to the speed and the user-defined parameter R_{outer}	210
8.3	The AOI is translated from P to P', according to the user's speed and trajectory.	211
8.4	d depends linearly on the user's speed	211
8.5	Area of interest for two user's contexts.	213
8.6	LT-REJA prototype architecture. Contextual information is incorporated with the contextual post-filtering approach.	214
8.7	LT-REJA prototype screenshots for various contextual situations. User's location is depicted with the green pin, and recommended restaurants with red pins. The adjusted AOI has been depicted in these figures with a blue circle ($R_{outer}=2\text{km}$).	214
8.8	General scheme of the context-aware group recommender of REJA	216
8.9	Architecture of the prototype for location-aware consensus-driven group recommendations.	219
8.10	Screens of the prototype for the login task.	220
8.11	Screens of the prototype for the group formation task.	221
8.12	Screenshots of prototype for the recommendation task.	222

List of Tables

2.1	Ratings table of a collaborative filtering recommender system . . .	46
2.2	Publications in recommender systems according to the recommendation domain and technique applied.	55
2.3	Notation used group recommender systems to refer to ratings with their respective interpretation in terms of the set of elements that they refer to.	58
2.4	Illustrative example for the classification of the ratings as possibly noisy.	93
3.1	Results for NRMSE of Mean, RMSMean, HGRM, and HGRM no-dup across various group sizes.	105
3.2	Results for $1 - NDCG$ of Mean, RMSMean, HGRM, and HGRM no-dup across various group sizes for a recommendation list of 5 items.	105
3.3	Results for ILS of Mean, RMSMean, HGRM, and HGRM no-dup across various group sizes for a recommendation list of 5 items. .	106
3.4	Results for $NRMSE \cdot \alpha + ILS \cdot (1 - \alpha)$, $\alpha = 0.75$	107
3.5	Results for $NRMSE \cdot \alpha + ILS \cdot (1 - \alpha)$, $\alpha = 0.50$	107
3.6	Results for $NRMSE \cdot \alpha + ILS \cdot (1 - \alpha)$, $\alpha = 0.25$	107
3.7	Computational complexity of proposal tasks.	109
3.8	Computational complexity of evaluated approaches tasks.	109

4.1	Evaluation results according to AUC. Larger values indicate better performance.	124
4.2	Evaluation results according to MAE. Smaller values indicate better performance.	124
4.3	Group recommendations generated by the baseline GRS and the consensus-based GRS	128
4.4	Computational complexity of proposal tasks.	129
4.5	Computational complexity of evaluated approaches tasks.	129
5.1	Ratings considered for the example of application of Pre-GROD. .	139
5.2	Similarity matrix extracted from the ratings in Table 5.1.	139
5.3	Weights matrix derived from similarity matrix	139
5.4	Parameters for the individual predictor	142
5.5	Main features of the datasets used in the experiments over the framework of group recommendation based on opinion dynamics. .	143
5.6	MAE for the evaluated GRSs in MovieLens 100k	144
5.7	RMSE for the evaluated GRSs in MovieLens 100k	144
5.8	MAE for the evaluated GRSs in MovieLens 1m	147
5.9	RMSE for the evaluated GRSs in MovieLens 1m	147
5.10	MAE for the evaluated GRSs in MovieLens 100k	148
5.11	RMSE for the evaluated GRSs in MovieLens 100k	148
5.12	MAE for the evaluated GRSs in MovieLens 1M	151
5.13	RMSE for the evaluated GRSs in MovieLens 1M	151
5.14	Computational complexity of proposal tasks.	151
5.15	Computational complexity of evaluated approaches tasks.	151
6.1	Illustrative example for the classification of the ratings as possibly noisy.	169

6.2	Corrected ratings R^* outputted by the NNM-GG correction applied over Table 6.1. The values corrected by NNM-GG are highlighted in bold. This rating database is R^* and it is used as input for NNM-LG to produce $R_{Ga,\bullet}^{**}$	171
6.3	Results of mean absolute error for the compared natural noise management approaches on recommendation aggregation group recommender systems.	174
6.4	Results of mean absolute error for the compared natural noise management approaches on rating aggregation group recommender systems.	176
6.5	Paired samples t-test p-values to compare each natural noise management technique with the baseline on MovieLens 100k dataset. .	178
6.6	Paired samples t-test p-values to compare each natural noise management technique with the baseline on Netflix Tiny dataset. . . .	180
6.7	Paired samples t-test p-values to compare NNM-GG with NNM-H.	181
7.1	Main stats of some StackExchange site datasets.	198
A.1	Similarity matrix a group of ten users that have no opinion leaders (rounded to two decimals).	271
A.2	Linear combination of the initial opinions for each opinion group (rounded to two decimals).	271
A.3	Decomposition of the group in subgroups without opinion leaders. The user added in each step is marked with the label of the network partition that it belongs to.	272

